

Title (en)
DOCUMENT SORTING METHOD, DOCUMENT SORTER, AND RECORDED MEDIUM ON WHICH DOCUMENT SORTING PROGRAM IS RECORDED

Title (de)
DOKUMENTENSORTIERVERFAHREN, DOKUMENTENSORTIERER UND AUFZEICHNUNGSMEDIUM AUF DEM DAS DOKUMENTENSORTIERPROGRAMM AUFGENOMMEN IST

Title (fr)
PROCEDE DE TRI DE DOCUMENTS, APPAREIL DE TRI DE DOCUMENTS ET SUPPORT ENREGISTRE SUR LEQUEL UN PROGRAMME DE TRI DE DOCUMENTS EST MEMORISE

Publication
EP 1124189 A1 20010816 (EN)

Application
EP 00931690 A 20000602

Priority
• JP 0003625 W 20000602
• JP 15849899 A 19990604
• JP 21250199 A 19990727

Abstract (en)
A document categorizing apparatus includes a sentence analyzer 12 for analyzing a plurality of documents to detect titles thereof; a feature element extractor 13 for extracting feature elements from the titles detected by the sentence analyzer 12 from the respective documents; feature table generating means 14 for generating a feature table representing the relationships between the feature elements extracted from the title and the documents including the feature elements; a document categorizing unit 15 for categorizing the documents into a plurality of clusters according to semantic similarity on the basis of the content of the feature table; a categorization result storage unit 16 for storing the clusters created by the document categorization unit 15; a cluster merging unit 2 for performing a cluster merging process upon the clusters stored in the categorization result storage unit 6; and an output control unit 31 for outputting the result of the cluster merging process to a display unit 32.

IPC 1-7
G06F 17/30

IPC 8 full level
G06F 17/27 (2006.01); **G06F 17/30** (2006.01)

CPC (source: EP US)
G06F 16/93 (2018.12 - EP US); **Y10S 707/99933** (2013.01 - US); **Y10S 707/99935** (2013.01 - US)

Cited by
GB2379051B; GB2412988A; GB2412988B; WO2016122512A1

Designated contracting state (EPC)
DE FR GB

DOCDB simple family (publication)
EP 1124189 A1 20010816; **EP 1124189 A4 20040721**; AU 4954200 A 20001228; US 7213205 B1 20070501; WO 0075810 A1 20001214

DOCDB simple family (application)
EP 00931690 A 20000602; AU 4954200 A 20000602; JP 0003625 W 20000602; US 76212600 A 20000602